

# Confusion Matrix

Ki Hyun Kim

[nlp.with.deep.learning@gmail.com](mailto:nlp.with.deep.learning@gmail.com)

# Accuracy is not Enough in Multi-class Classification

- 정확도(accuracy)를 통해 모델의 성능을 순서대로 나타낼 수 있음
- 하지만 하나의 숫자로 나타낸 만큼 내부의 자세한 성능은 알 수 없다.
  - 만약 test-set의 class 분포가 불균등(imbalance)하다면?
  - 성능을 개선(e.g. 제품화 & 배포)할 때, 부족한 부분을 쉽게 파악할 수 없다.
- 각 class별 성능을 파악할 수 있을까?

# Confusion Matrix in MNIST Classification

- 어떤 class 사이의 추론 성능이 떨어지는지 쉽게 확인 & 대처 가능

	pred_0	pred_1	pred_2	pred_3	pred_4	pred_5	pred_6	pred_7	pred_8	pred_9
true_0	972	0	0	0	1	0	1	0	4	2
true_1	0	1122	1	3	0	1	2	1	5	0
true_2	5	0	1016	3	1	0	0	4	3	0
true_3	2	0	6	975	0	9	0	7	4	7
true_4	4	0	4	0	953	0	6	2	1	12
true_5	6	0	0	4	2	863	4	1	6	6
true_6	4	2	0	0	9	7	932	0	4	0
true_7	1	1	12	1	0	0	0	1006	2	5
true_8	6	0	7	3	3	1	2	5	946	1
true_9	6	2	0	5	12	1	0	5	6	972